

# Cell Population Dynamics: Its Relationship with Finite State Markov Chain and its Asymptotic Behavior

Da-Quan Jiang\*

Yue Wang<sup>†</sup>Da Zhou<sup>‡</sup>

## Abstract

We consider the cell population dynamics with  $n$  different phenotypes. Cells in one phenotype can produce cells in other phenotypes through conversions or asymmetric divisions. Both the Markov branching process model and the ordinary differential equation (ODE) system model are presented, and exploited to investigate the dynamics of the phenotypic proportions. [11] observed that with different initial population states, the proportions of different phenotypes will always tend to certain constants (“phenotypic equilibrium”). In the ODE system model, they gave a mathematical explanation through assuming the phenotypic proportions satisfy the Kolmogorov forward equations of an  $n$ -state Markov chain. We give a sufficient and necessary condition under which this assumption is valid. We also prove the “phenotypic equilibrium” without such assumption. In the Markov branching process model, more generally, we show the stochastic explanation of “phenotypic equilibrium” through improving a limit theorem in [14], which may be of theoretical interests. As an application, we will give sufficient and necessary conditions under which the proportion of one phenotype tends to 0 (die out) or 1 (dominate). We also extend our results to non-Markov cases.

**KEY WORDS:** population dynamics, Markov chains, asymptotic behavior, branching processes, phenotypic equilibrium

**2010 Mathematics Subject Classification:** 60J85, 92D25, 34D05

## 1 Introduction

With the same genetic background, cell population may have different cellular phenotypes. This has been one of the major topics in the research of cell population dynamics [1, 18]. Very recently much attention has been paid to the stochastic conversions between different phenotypes [6, 11]. For example, we know that cancer stem cells can give rise to cancer

---

\*LMAM, School of Mathematical Sciences & Center for Statistical Science, Peking University, Beijing 100871, P.R. China. Email address: jiangdq@math.pku.edu.cn

<sup>†</sup>Department of Applied Mathematics, University of Washington, Seattle, WA 98195, USA. Email address: yuewang@uw.edu

<sup>‡</sup>School of Mathematical Sciences, Xiamen University, Xiamen 361005, P.R. China. Email address: zhouda@xmu.edu.cn

non-stem cells, but cancer non-stem cells can also transform back to cancer stem cells [31]. Generally, we can use a branching process (stochastic model) [13, 17, 26, 27, 28] or an ODE system (deterministic model) [20] to describe the dynamics of such cell population with multiple phenotypes. However, in many experimental settings, it is difficult or even impossible to count the total cell population [4, 27, 28]. Thus in the last fifty years, people began to consider the proportions of cell individuals with distinct phenotypes instead of the absolute numbers of cells of various phenotypes [13].

In the experiments on breast cancer cell lines, [11] found that the proportion of each phenotype will always tend to a certain constant regardless of the initial population states (“phenotypic equilibrium”). They assumed that the evolution of the phenotypic proportions satisfies an  $n$ -state Markov chain, and used the ergodicity of the Markov chain to explain this phenomenon [11]. However, we find that this assumption is not always valid. We determine the condition under which this assumption is valid. Furthermore, we try to remove this assumption and explain the experimental phenomenon in [11] under more general context.

In deterministic model (ODE system), we only consider the average behavior of cell population dynamics. However, using stochastic model (branching processes), we can study the trajectory behavior. We prove that the proportions will converge not only on average, but also almost surely.

In the theory of multi-type branching processes, people have observed similar proportion convergence phenomenon and proved such phenomenon in several limit theorems under different conditions [16, 3, 14]. Those are possible ways to explain “phenotypic equilibrium”, but those required conditions may not be satisfied in experiments. Thus we improve those limit theorems by dropping redundant conditions. We will see that the conditions we need are all biologically reasonable. Therefore, we give a stochastic explanation of “phenotypic equilibrium”. This result may also be of interests to probabilists.

Generally we only consider Markov branching processes, but sometimes the biological process is not memoryless, thus we need to consider non-Markov branching processes. We show that under some conditions, the non-Markov branching processes can be transformed into Markov branching processes. Using this trick, we demonstrate similar results for non-Markov branching processes.

In Section 2, we will give the mathematical description of our models, which is based on [17] and [14]. In Section 3, we will give a sufficient and necessary condition under which the assumption in [11] is valid. In Section 4, we will prove that under some mild conditions, the “phenotypic equilibrium” phenomenon will always happen in the Markov branching process model. Specifically, we will improve a limit theorem about proportion convergence in multi-type branching processes. In Section 5, as an application of our conclusions, we will investigate under what conditions one of the phenotypes will die out or dominate. In Section 6, we will show that the above conclusions are still valid in more general cases.

## 2 Model description

**Stochastic Model.** Assume that the population of cells have  $n$  phenotypes:  $X_1, X_2, \dots, X_n$ . Assume that all the cells evolve independently. ( During the exponential growth period, this assumption is almost true [32]. ) We can present the generalized cell divisions, death and phenotypic conversions as the following reaction form:

$$X_i \xrightarrow{\alpha_i} d_{i1}X_1 + d_{i2}X_2 + \dots + d_{in}X_n.$$

It means that for an  $X_i$  cell, it will live an exponential time (we will consider non-exponential lifetime in Section 6) with expectation  $1/\alpha_i$  and turn into  $d_{i1}$   $X_1$  cells,  $d_{i2}$   $X_2$  cells,  $\dots$ ,  $d_{in}$   $X_n$  cells, where  $d_{i1}, d_{i2}, \dots, d_{in}$  are random variables taking nonnegative integer values.  $d_{i1}, d_{i2}, \dots, d_{in}$  are not necessarily independent, but they are assumed to be independent of the exponential reaction time. In fact this is a continuous-time branching process with state space  $(\mathbb{Z}^*)^n$ , each component of which represents the population of a phenotype. It is also called the generalized Pólya urn. For example, if one  $X_1$  cell splits symmetrically, the process will move from the state  $(s_1, s_2, \dots, s_n)$  to the state  $(s_1 + 1, s_2, \dots, s_n)$ . We require that  $\mathbb{E}d_{ij}^2 < \infty, \forall i, j$ . (In experiments,  $d_{ij}$  is bounded, thus  $\mathbb{E}d_{ij}^2 < \infty$  is always true.) Then this process will not explode in finite time with probability one [3, Section V.7.1, (3)–(4)].

**Deterministic Model.** Now we consider the expectation of the populations of the  $n$  phenotypes at time  $t$ ,  $(x_1(t), x_2(t), \dots, x_n(t))$ . Based on [3, Section V.7.2, (5)–(9)], we have the deterministic model, namely the following ODE system:

$$\begin{cases} \frac{dx_1}{dt} = a_{1,1}x_1 + a_{1,2}x_2 + \dots + a_{1,n}x_n, \\ \frac{dx_2}{dt} = a_{2,1}x_1 + a_{2,2}x_2 + \dots + a_{2,n}x_n, \\ \vdots \\ \frac{dx_n}{dt} = a_{n,1}x_1 + a_{n,2}x_2 + \dots + a_{n,n}x_n. \end{cases} \quad (1)$$

where  $a_{i,i} = \alpha_i(\mathbb{E}d_{ii} - 1) \geq -\alpha_i$ ,  $a_{i,j} = \alpha_j\mathbb{E}d_{ji} \geq 0$  ( $i \neq j$ ). Define  $\mathbf{A} = \begin{bmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \dots & a_{n,n} \end{bmatrix}$ ,

the coefficient matrix of (1).

## 3 The relation between the $n$ -state Markov chain and the deterministic model

In this section, we will discuss when the deterministic model can be equivalently captured by the Kolmogorov forward equations of an  $n$ -state Markov chain. Thus we can verify when the assumption in [11] is valid.

First we consider the proportions of each expected subpopulation  $x_1(t), x_2(t), \dots, x_n(t)$  in (1) among the expected whole population  $x_1(t) + \dots + x_n(t)$ .

Define

$$p_1(t) = \frac{x_1(t)}{x_1(t) + x_2(t) + \cdots + x_n(t)}, \cdots, p_n(t) = \frac{x_n(t)}{x_1(t) + x_2(t) + \cdots + x_n(t)}.$$

Using  $p_n(t) = 1 - \sum_{i=1}^{n-1} p_i(t)$ , we can get the differential equations of  $p_1(t), \cdots, p_{n-1}(t)$  from (1):

$$\begin{cases} \frac{dp_1}{dt} = \sum_{i=1}^{n-1} A_i p_1 p_i + \sum_{i=1}^{n-1} B_{1,i} p_i + a_{1,n}, \\ \frac{dp_2}{dt} = \sum_{i=1}^{n-1} A_i p_2 p_i + \sum_{i=1}^{n-1} B_{2,i} p_i + a_{2,n}, \\ \vdots \\ \frac{dp_{n-1}}{dt} = \sum_{i=1}^{n-1} A_i p_{n-1} p_i + \sum_{i=1}^{n-1} B_{n-1,i} p_i + a_{n-1,n}. \end{cases} \quad (2)$$

where  $A_i = -\sum_{j=1}^n a_{j,i} + \sum_{j=1}^n a_{j,n}$ ,  $B_{i,i} = a_{i,i} - a_{i,n} - \sum_{i=1}^n a_{i,n}$ ,  $B_{i,j} = a_{i,j} - a_{i,n} (i \neq j)$ .

Now consider an  $n$ -state continuous-time Markov chain with Q-matrix  $\{q_{i,j}\}$ . We can describe it by the Kolmogorov forward equations:

$$\begin{cases} \frac{dP_1(t)}{dt} = q_{1,1}P_1(t) + q_{2,1}P_2(t) + \cdots + q_{n,1}P_n(t), \\ \frac{dP_2(t)}{dt} = q_{1,2}P_1(t) + q_{2,2}P_2(t) + \cdots + q_{n,2}P_n(t), \\ \vdots \\ \frac{dP_n(t)}{dt} = q_{1,n}P_1(t) + q_{2,n}P_2(t) + \cdots + q_{n,n}P_n(t). \end{cases} \quad (3)$$

where  $P_i(t)$  is the probability of the Markov chain being in state  $i$  at time  $t$ . Using  $\sum_{i=1}^n P_i(t) = 1$  to remove  $P_n(t)$ , we can rewrite (3) as:

$$\begin{cases} \frac{dP_1(t)}{dt} = (q_{1,1} - q_{n,1})P_1(t) + (q_{2,1} - q_{n,1})P_2(t) + \cdots + (q_{n-1,1} - q_{n,1})P_{n-1}(t) + q_{n,1}, \\ \frac{dP_2(t)}{dt} = (q_{1,2} - q_{n,2})P_1(t) + (q_{2,2} - q_{n,2})P_2(t) + \cdots + (q_{n-1,2} - q_{n,2})P_{n-1}(t) + q_{n,2}, \\ \vdots \\ \frac{dP_{n-1}(t)}{dt} = (q_{1,n-1} - q_{n,n-1})P_1(t) + \cdots + (q_{n-1,n-1} - q_{n,n-1})P_{n-1}(t) + q_{n,n-1}. \end{cases} \quad (4)$$

In order that (2) has the same form of (4), all second-order coefficients  $A_i$  in (2) should be 0, namely

$$K := \sum_{i=1}^n a_{i,1} = \sum_{i=1}^n a_{i,2} = \cdots = \sum_{i=1}^n a_{i,n-1} = \sum_{i=1}^n a_{i,n}. \quad (5)$$

If so, we can rewrite (2) as:

$$\begin{cases} \frac{dp_1(t)}{dt} = (a_{1,1} - K)p_1(t) + a_{1,2}p_2(t) + \cdots + a_{1,n}p_n(t), \\ \frac{dp_2(t)}{dt} = a_{2,1}p_1(t) + (a_{2,2} - K)p_2(t) + \cdots + a_{2,n}p_n(t), \\ \vdots \\ \frac{dp_n(t)}{dt} = a_{n,1}p_1(t) + a_{n,2}p_2(t) + \cdots + (a_{n,n} - K)p_n(t). \end{cases} \quad (6)$$

Notice (5) and that  $a_{i,j} (i \neq j)$  is nonnegative, (6) has the same form as (3). Thus we have

**Theorem 1.** *Equation (5) is the sufficient and necessary condition for that the proportions of different phenotypes in the deterministic model (1) satisfy the Kolmogorov forward equations of an  $n$ -state Markov chain.*

With Theorem 1 we can analyze the asymptotic behavior of the deterministic model of population dynamics exploiting the  $n$ -state Markov chain. From the Markov chain theory [21], we know that if  $\mathbf{A}$  is irreducible, then the solution of (6) will converge to the unique invariant distribution, no matter what the initial values are. This is just the mathematical basis of [11]. It has been reported that the condition in Theorem 1 is satisfied not only in the breast cancer cell lines in Gupta et al's experiments, but also in colon cancer cell lines [29, 31].

## 4 Asymptotic behavior in general cases

In general cases, (5) is not satisfied since different phenotypes may differ in cell cycling time [23, 9], then the  $n$ -state Markov chain simplification is invalid. Thus we need other methods to study the asymptotic behavior of the population dynamics. In this section, we will prove that under some mild conditions, the proportions of different phenotypes will tend to some constants regardless of initial population states.

From Perron-Frobenius theorem [24, 15], we know that  $\mathbf{A}$  has a real eigenvalue  $\lambda_1$  (called Perron eigenvalue), such that for any eigenvalue  $\mu \neq \lambda_1$ ,  $\text{Re } \mu < \lambda_1$ .  $\lambda_1$  has a right eigenvector  $\mathbf{u} = (u_1, u_2, \dots, u_n)$  (called Perron eigenvector), satisfying  $u_i \geq 0, \forall i$  and  $\sum_{i=1}^n u_i = 1$ . When  $\lambda_1$  is simple, such  $\mathbf{u}$  is unique. We know that the set of all  $n$ -order real square matrices with repeated eigenvalue has measure 0 (as a subset of  $\mathbb{R}^{n^2}$ ) [30]. Thus it is reasonable to assume that  $\lambda_1$  is simple.

### 4.1 deterministic model

We have proved the following theorem in Appendix B of [30].

**Theorem 2.** *Assume that  $\lambda_1$  is simple. Starting from any initial value except for the point in some zero-measure set, we have  $(x_1(t), x_2(t), \dots, x_n(t)) / \exp(\lambda_1 t) \rightarrow c\mathbf{u}$  as  $t \rightarrow \infty$ , where  $c > 0$  is a constant. In this case, the solution of (2) will tend to  $\mathbf{u}$  as  $t \rightarrow \infty$ . Thus (2) has one and only one stable fixed point  $\mathbf{u}$  and no stable limit cycle.*

This gives a satisfactory deterministic explanation of the phenotypic equilibrium phenomenon reported in [11].

**Remark 1.** *If  $\lambda_1$  is not simple, then the convergence result may not hold. Consider  $\mathbf{A}$  with  $a_{i,j} = 0, \forall i, j$ . Here  $\lambda_1 = 0$  is not simple, and the system will never move. Convergence to a common point will never occur.*

## 4.2 stochastic model

Let  $\mathbf{X}^*(t) = (X_1^*(t), X_2^*(t), \dots, X_n^*(t))$  be the population of  $n$  phenotypes at time  $t$ .

Since 1960s, probabilists proved that  $(X_1^*(t), X_2^*(t), \dots, X_n^*(t))/e^{\lambda_1 t} \rightarrow W\mathbf{u}$  under different conditions, where  $W$  is a nonnegative random variable. In [2], [3] and [25], it is required that  $\lambda_1 > 0$  and  $\mathbf{A}$  is irreducible (this implies  $\lambda_1$  is simple). In [3] it is proved that  $W = 0$  or  $W > 0$  according to whether the population will become extinct. In [16], it is required that the branching process is discrete in time. In [27, 28] it is required that the initial population tends to infinity. [14] requires that  $\lambda_1 > 0$ ,  $\lambda_1$  is simple, and assumes a special condition about communicating classes structure (see Remark 2). So far [14] is the best result about this problem. Based on [14] and [3], we will prove the convergence theorem without Janson's last assumption (Theorem 3). We can see the benefit of this improvement in Section 5.

### 4.2.1 preliminaries

In this section, we assume that  $\lambda_1$  is simple and positive.  $\lambda_1 > 0$  means that the total cell population is increasing.

Sometimes, the transformation from one phenotype to another phenotype is not reversible. For example, a mature human red blood cell (which loses its nucleus) cannot transform back to a zygote. Thus we need to classify phenotypes according to communicating behaviors. In mathematical language, we need to study communicating classes of  $\mathbf{A}$  when  $\mathbf{A}$  is reducible.

We can divide the  $n$  phenotypes into several communicating classes according to  $\mathbf{A}$ . Then we can order the classes and rearrange the phenotypes suitably to make  $\mathbf{A}$  block-triangular. (Each diagonal block corresponds to a communicating class.) Thus the eigenvalues of  $\mathbf{A}$  consist of all eigenvalues of diagonal blocks. Every eigenvalue corresponds to a diagonal block, and then corresponds to a communicating class. (See [14] and [16] for details.)

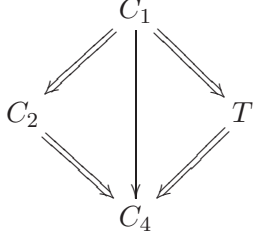
Denote the communicating class corresponding to the Perron eigenvalue  $\lambda_1$  by  $T$ .

For example, consider matrix  $\mathbf{A} = \begin{bmatrix} \mathbf{D}_1 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{X} & \mathbf{D}_2 & \mathbf{0} & \mathbf{0} \\ \mathbf{X} & \mathbf{0} & \mathbf{D}_3 & \mathbf{0} \\ \mathbf{0} & \mathbf{X} & \mathbf{X} & \mathbf{D}_4 \end{bmatrix}$ , where each  $\mathbf{X}$  represents a

different nonnegative matrix (not  $\mathbf{0}$ ). Assume that  $\mathbf{D}_3$  has the Perron eigenvalue  $\lambda_1$ , then  $\mathbf{D}_3$  corresponds to the communicating class  $T$ . Denote the other three communicating classes by  $C_1, C_2, C_4$ .

For two communicating classes  $C_i$  and  $C_j$ , we write  $C_i \Rightarrow C_j$  if there exist phenotype  $X_{k_i} \in C_i$  and  $X_{k_j} \in C_j$  such that  $a_{k_j, k_i} > 0$ . For two communicating classes  $C$  and  $D$ , we write  $C \rightarrow D$  if there exist communicating classes  $C = C_1, C_2, \dots, C_m = D$  such that  $C_i \Rightarrow C_{i+1}, \forall 1 \leq i < m$ . Stipulate that  $C_i \Rightarrow C_i$  and  $C_i \rightarrow C_i$ .

Then we can illustrate the communicating classes in the example above as



For a communicating class  $C$ , define  $\hat{C} = \{X_i | X_i \in C_j, C_j \rightarrow C\}$ . In other words,  $\hat{C}$  is the set of all phenotypes that can produce (directly or indirectly) phenotypes in  $C$ . In the example above  $\hat{T} = C_1 \cup T$ .

For a communicating class  $C$ , define  $\bar{C} = \{X_i | X_i \in C_j, C \rightarrow C_j\}$ . In other words,  $\bar{C}$  is the set of all phenotypes that can be produced (directly or indirectly) by phenotypes in  $C$ . In the example above  $\bar{T} = T \cup C_4$ .

For the Markov branching process  $\mathbf{X}^*(\cdot)$ , we say that a cell  $Y$  with phenotype in  $\hat{T}$  becomes “*essentially extinct*” if at some time no cell of any phenotypes in  $\hat{T}$  is  $Y$  or its descendants. In other words,  $Y$  and its descendants become extinct inside  $\hat{T}$ . We say that a trajectory of the branching process  $\mathbf{X}^*(\cdot)$  becomes “*essentially extinct*” if at some time no cell of any phenotypes in  $\hat{T}$  remains. This means that we can never get a cell with phenotypes in  $T$  any more. If so, we cannot have the desired convergence property. Let the branching process  $\mathbf{X}^*(\cdot)$  start at any initial population  $\mathbf{X}^*(0)$  as long as it has some cells with phenotypes in  $\hat{T}$ .

Let  $P_i^*(t) = X_i^*(t) / \sum_{i=1}^n X_i^*(t)$  be the proportion of phenotype  $X_i$ , as long as the denominator is not zero.

#### 4.2.2 results and proofs

We now state the main result of this paper and then give the proof of it.

**Theorem 3.** *Assume that  $\lambda_1$  is simple and positive. Conditioned on essential non-extinction, we have almost surely  $(P_1^*(t), P_2^*(t), \dots, P_n^*(t)) \rightarrow \vec{u} = (u_1, u_2, \dots, u_n)$  as  $t \rightarrow \infty$ .*

**Lemma 1** (Lemma 9.8 in [14]). *Assume that  $\lambda_1$  is simple and positive. Then we have almost surely  $e^{-\lambda_1 t} \mathbf{X}^*(t) \rightarrow W \mathbf{u}$  as  $t \rightarrow \infty$ , where  $W$  is a nonnegative random variable, and  $\mathbb{P}(W > 0) > 0$ .*

**Lemma 2** (Lemma 9.7 (ii) and (iii) in [14], originated from Theorem V.7.2 in [3]). *Assume that  $\lambda_1$  is simple and positive, and  $\bar{T}$  contains all phenotypes, then  $W = 0$  if and only if the branching process becomes essentially extinct almost surely.*

**Remark 2.** [14, Section 2] has six fundamental assumptions (A1)-(A6). Assumptions (A1)-(A5) have been satisfied in this paper (regarding (A5) as “the process is not essentially extinct at time 0”). Assumption (A6) “ $\bar{T}$  contains all phenotypes” is only used in Lemma 2. We will remove this assumption in Lemma 5.

The following lemma is a modification of the second Borel-Cantelli lemma. We base our proof on Theorem 2.3.6 in [7].

**Lemma 3.** *Consider events  $B_1, B_2, \dots, B_n, \dots$ . If for any positive integers  $m < n$ , we have  $\mathbb{P}(\cap_{i=m+1}^n B_i^c) \leq (1 - \epsilon)^{n-m}$ , where  $0 < \epsilon \leq 1$ , then  $\mathbb{P}(\limsup_{n \rightarrow \infty} B_n) = 1$ . In other words, almost surely  $\{B_n : n \geq 1\}$  will happen infinitely often.*

*Proof.* Let  $0 < M < N < \infty$ .  $\mathbb{P}(\cap_{i=M+1}^N B_i^c) \leq (1 - \epsilon)^{N-M} \rightarrow 0$  as  $N \rightarrow \infty$ . So  $\mathbb{P}(\cup_{i=M+1}^\infty B_i) = 1$  for all  $M$ , and since  $\cup_{i=M+1}^\infty B_i \downarrow \limsup_{n \rightarrow \infty} B_n$  it follows that  $\mathbb{P}(\limsup_{n \rightarrow \infty} B_n) = 1$ .  $\square$

**Lemma 4.** *For almost every essentially non-extinct trajectory (according to Lemma 1, the set of such trajectories has positive probability), we can find an essentially non-extinct cell with phenotype in  $T$  within finite time. If we can find such cell at time  $t$ , then we can find such cell at any time  $\tau > t$ .*

*Proof.* If at some time  $t$  all cells with phenotypes in  $\hat{T} \setminus T$  die out, then at least one of the remaining cells with phenotypes in  $T$  is not essentially extinct.

Otherwise, at each time  $t = k$  ( $k \in \mathbb{Z}^+$ ), there exists one cell  $E_k$  with phenotype in  $\hat{T} \setminus T$ . (For different  $k$ ,  $E_k$  may be the same cell.) Let  $B_k$  ( $k \in \mathbb{Z}^+$ ) be the event that during the time interval  $[k, k+1)$ , the cell  $E_k$  produces (directly or indirectly) at least one cell with phenotype in  $T$ .

If  $B_k$  happens, choose one such cell with phenotype in  $T$  and put it in a special set  $S$ . Consider any two cells  $F$  and  $G$  in  $S$ , and assume  $F$  is produced in the time interval  $[i, i+1)$ ,  $G$  is produced in the time interval  $[j, j+1)$ , and  $i < j$ , where  $i, j \in \mathbb{Z}^+$ . Then  $E_j$  is the ancestor of  $G$ . Since  $E_j$  has phenotype in  $\hat{T} \setminus T$ , and  $F$  has phenotype in  $T$ ,  $F$  cannot be the ancestor of  $E_j$ . Since  $E_j$  is still alive at time  $t = j$ , when  $F$  has been produced,  $E_j$  cannot be the ancestor of  $F$ . Thus  $F$  cannot be the ancestor of  $G$ . Since  $G$  is produced after  $F$ ,  $G$  cannot be the ancestor of  $F$ . In sum, one cell in  $S$  cannot be the ancestor of another cell in  $S$ . Thus all cells in  $S$  are independent.

Consider two phenotypes  $X_i$  and  $X_j$ , and assume a cell with phenotype  $X_i$  can produce a cell with phenotype  $X_j$  directly, namely  $\mathbb{P}(d_{ij} > 0) > 0$ . Because of Markov property, within a time span of  $1/n$ , the probability for a cell with phenotype  $X_i$  to produce a cell with phenotype  $X_j$  directly is  $\eta_{ij} = [1 - \exp(-\alpha_i/n)]\mathbb{P}(d_{ij} > 0) > 0$ . Let  $\eta = \min_{i,j} \{\eta_{i,j} : \mathbb{P}(d_{ij} > 0) > 0\}$ . For a cell with phenotype in  $\hat{T} \setminus T$ , it can produce a cell with phenotype in  $T$  within  $n$  steps. Thus the probability of  $B_k$  is no less than  $\eta^n$ , regardless of what happens before time  $t = k$ .

Now we can use Lemma 3 with  $\epsilon = \eta^n$ , and there will be an infinite number of cells in  $S$ , except for a zero-measure set of trajectories. According to Lemma 1, the probability for one cell in  $S$  to become essentially extinct is less than 1, thus the probability for all cells in  $S$  to become essentially extinct is 0, and at least one cell in  $S$  is not essentially extinct, except for a zero-measure set of trajectories.  $\square$

**Lemma 5.** *Assume that  $\lambda_1$  is simple and positive, then  $W = 0$  if and only if the branching process becomes essentially extinct almost surely.*



*Proof.*  $\Leftarrow$ : For a trajectory  $X^*(\cdot)$  outside the zero-measure exclusion set of Lemma 1, assume that at some time  $\tau \geq 0$  (dependent on the trajectory),  $X_i^*(\tau) = 0$  for all  $i \in \hat{T}$ . For any  $j \in T$ ,  $0 = \lim_{t \rightarrow \infty} e^{-\lambda_1 t} X_j^*(t) = W u_j$ . From Lemma 7,  $u_j > 0$ . Thus  $W = 0$  almost surely.

$\Rightarrow$ : Assume that  $\mathbb{P}(W = 0 \text{ \& the trajectory is not essentially extinct}) = P_0 > 0$ . According to Lemma 4, we can find time  $t_0 > 0$  large enough such that  $\mathbb{P}(W = 0 \text{ \& the trajectory is not essentially extinct \& there exists an essentially non-extinct cell with phenotype in } T \text{ at time } t_0) \geq P_0/2 > 0$ . On this set, only consider this essentially non-extinct cell and its descendants from time  $t \geq t_0$ , then the population is restricted on  $\bar{T}$  and we can use Lemma 2. Now we have  $W > 0$  except for a zero-measure set of trajectories, which is a contradiction.  $\square$

From Lemma 1 and Lemma 5, we can obtain Theorem 3.

**Remark 3.** *The assumption of  $\lambda_1 > 0$  is necessary. If  $\lambda_1 < 0$ , then from Theorem 2, the expected populations decay to 0. Therefore this process will become extinct almost surely. For  $\lambda_1 = 0$ , consider a special case that phenotype  $X_1$  can only transform to phenotype  $X_2$  and vice versa. Starting from one cell with phenotype  $X_1$ ,  $(P_1^*, P_2^*)$  will jump between  $(1, 0)$  and  $(0, 1)$ . This process will not become essentially extinct, and the proportions will not tend to constants [14].*

For Gupta et al's experiment, the initial cell population is very large in cancer cell lines, thus the probability of essential extinction is quite small. Therefore, the proportions will almost always tend to the same constants. This gives a satisfactory stochastic explanation of the phenotypic equilibrium phenomenon reported in [11].

## 5 When will one proportion tend to 0 or 1?

In population dynamics, we are also concerned about when one phenotype dies out or dominates. In terms of the notations in this paper, we need to consider when  $P_i^*(t) \rightarrow 0$  or  $P_i^*(t) \rightarrow 1$  as  $t \rightarrow \infty$ .

In this section, we will still assume that the Perron eigenvalue  $\lambda_1$  of  $\mathbf{A}$  is simple and positive. Then from Theorem 3, we have  $(P_1^*(t), P_2^*(t), \dots, P_n^*(t)) \rightarrow \mathbf{u} = (u_1, u_2, \dots, u_n)$  almost surely in the stochastic model. Then we need to study the properties of  $\mathbf{u}$ .

**Lemma 6.** *Assume that  $\lambda_1$  is simple. If for some  $i \neq j$ ,  $a_{i,j} > 0$  in (1), then  $u_j > 0 \Rightarrow u_i > 0$ .*

*Proof.* Without loss of generality, let  $i = 1, j = 2$ . Assume  $u_1 = 0, u_2 > 0$ . Let  $(p_1, p_2, \dots, p_n) = \mathbf{u} = (u_1, u_2, \dots, u_n)$  in the first equation of (2). Since  $a_{1,n} = a_{1,n} \sum_{k=2}^n u_k$ ,  $a_{1,2} > 0$  and  $u_2 > 0$ , the equation becomes  $dp_1/dt = \sum_{k=2}^n a_{1,k} u_k > 0$ . However  $\mathbf{u}$  is a fixed point of (2) according to Theorem 2, thus we should have  $dp_1/dt = 0$ , which is a contradiction.  $\square$

**Lemma 7.** Assume that  $\lambda_1$  is simple. Then  $u_i > 0 \iff X_i \in \bar{T}$ .

*Proof.* Apply the Perron-Frobenius theorem to  $\mathbf{A}_{\bar{T}}$ , the restriction of  $\mathbf{A}$  on  $\bar{T}$ , and let  $\mathbf{w}$  be its Perron eigenvector.  $\mathbf{w}_T$ , the restriction of  $\mathbf{w}$  on  $T$  cannot be  $\mathbf{0}$ , otherwise  $\lambda_1$  is an eigenvalue of  $\mathbf{A}_{\bar{T} \setminus T}$ , a contradiction. From Lemma 6 we know that  $\vec{w}$  is positive. Set  $u_i = w_i$  if  $X_i \in \bar{T}$ , and  $u_j = 0$  if  $X_j \notin \bar{T}$ , then  $\mathbf{u}$  is the Perron eigenvector of  $\mathbf{A}$ . Thus  $u_i > 0 \iff X_i \in \bar{T}$ . □

Now we can get the following theorems from Lemma 7.

**Theorem 4.**  $P_i^*(t) \rightarrow 0 \iff X_i \notin \bar{T}$ .

**Theorem 5.**  $P_i^*(t) \rightarrow 1 \iff \bar{T} = T = \{X_i\}$ .

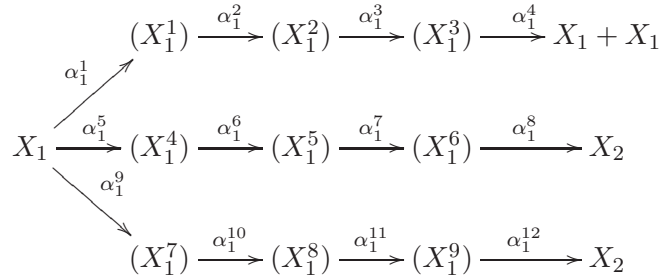
**Remark 4.** From Theorem 4 we can see that the sufficient and necessary condition under which no phenotype dies out, namely  $\forall i, P_i^*(t) \rightarrow 0$ , is that  $\bar{T}$  contains all phenotypes. This is just Janson's last assumption.

**Remark 5.** If we find that  $P_i^*(t) \rightarrow 0$ ,  $P_j^*(t) \rightarrow 0$  in an experiment, then we know that the phenotype  $X_j$  will never transform to  $X_i$  in any way. If we find that  $P_i^*(t) \rightarrow 1$ , then we know that the phenotype  $X_i$  will never transform to any other phenotypes.

## 6 Model generalization: non-exponential lifetime

In the previous sections, we assumed that the lifetime of a cell is exponentially distributed and independent of the type and number of its descendants. However, in real biological system, the lifetime distribution should be more like lognormal, gamma, Weibull, or exponentially modified Gaussian distribution [12, 10]. Furthermore, the time needed for division and conversion have different distributions [10]. In this way the branching process is no longer Markovian.

We can use the “device of stages” method to approximate a non-exponential random variable with several exponential random variables [5]. This indicates that through adding supplementary sub-phenotypes, we can simulate a non-Markov branching process with a Markov branching process. See the example below:



Here  $(X_1^1), \dots, (X_1^9)$  are supplementary sub-phenotypes, which are not distinguishable from  $X_1$  in experiments. An  $X_1$  cell has probability  $\alpha_1^1/(\alpha_1^1 + \alpha_1^5 + \alpha_1^9)$  to divide into  $X_1 + X_1$ , and probability  $(\alpha_1^5 + \alpha_1^9)/(\alpha_1^1 + \alpha_1^5 + \alpha_1^9)$  to convert into  $X_2$ . Here we set  $\alpha_1^1, \alpha_1^5$ , and  $\alpha_1^9$  to be large enough while keeping their proportions, so that the time needed for the first step is ignorable (exponential random variable with expectation  $1/(\alpha_1^1 + \alpha_1^5 + \alpha_1^9)$ ).

Now the time distribution for division  $X_1 \rightarrow X_1 + X_1$  is approximately  $Ex(\alpha_1^2) * Ex(\alpha_1^3) * Ex(\alpha_1^4)$ , where  $Ex(\alpha)$  is the density function of exponential random variable with parameter  $\alpha$ , and  $*$  means convolution. Similarly, the time distribution for conversion  $X_1 \rightarrow X_2$  is approximately  $\frac{\alpha_1^5}{\alpha_1^5 + \alpha_1^9} Ex(\alpha_1^6) * Ex(\alpha_1^7) * Ex(\alpha_1^8) + \frac{\alpha_1^9}{\alpha_1^5 + \alpha_1^9} Ex(\alpha_1^{10}) * Ex(\alpha_1^{11}) * Ex(\alpha_1^{12})$ .

According to [5], any non-negative random variable can be approximated by such combination of convolutions of exponential random variables. Thus we can simulate such non-Markov branching processes to any precision with Markov branching processes. Here the lifetime of a cell can be non-exponential, and the lifetime of a cell can depend on the type and number of its descendants.

Now we can apply Theorem 3 to those sub-phenotypes. The proportion of each sub-phenotype converges to a constant. Thus the proportion of each phenotype (including all its sub-phenotypes) converges to a constant. This proves the “phenotypic equilibrium” phenomenon in a more realistic stochastic model. In addition, the conclusions in Section 5 are still valid.

**Remark 6.** *The proportion convergence theorem for non-Markov (age-dependent) branching processes can be proved directly, but under stronger conditions [19].*

## 7 Conclusion

We have presented a unified stochastic model for the population dynamics with cellular phenotypic conversions. We have given the sufficient and necessary condition under which the dynamical behavior of our model can be described by an  $n$ -state Markov chain. In general case, we have proved that the proportions of different phenotypes will tend to constants regardless of their initial values, and we have investigated the sufficient and necessary conditions under which one phenotype will die out or dominate. We also extend our model to non-Markov case while keeping the above conclusions valid. In this way we have rigorously explained the experimental phenomenon in [11].

As remarked in Section 4.2, we improve a limit theorem in branching processes, which may be of theoretical interests.

Since the phenotypic conversions have been reported in various cellular systems, such as *E.coli* [22] and cancer cells [8, 29], we hope that our model here could be applied as a general framework in the study of multi-phenotypic populations of cells.

This research has some improvement spaces. First, we assume that the branching process is time homogeneous, namely the birth and death rates keep the same for all time. However, as time goes on, the cell density increases, and the birth and death rates should change [32]. Thus a possible improvement is to have time-dependent or density-dependent  $d_{ij}$ . Second,

we only prove the convergence for  $t \rightarrow \infty$ , but in experiments we only have finite observation time. Thus it is meaningful to estimate the convergence rate.

## Acknowledgements

We would like to thank Professor Min-Ping Qian, Da-Yue Chen and Svante Janson for helpful advice and discussions. D.-Q. J. acknowledges the support of National Natural Science Foundation of China 11271029 and National Natural Science Foundation of China 11171024. D. Z. acknowledges the support of National Natural Science Foundation of China 11401499 and the Natural Science Foundation of Fujian Province of China (No. 2015J05016).

## References

- [1] S. J. Altschuler and L. F. Wu. Cellular heterogeneity: do differences make a difference? *Cell*, 141(4):559–563, 2010.
- [2] K. B. Athreya. Some results on multitype continuous time Markov branching processes. *Ann. Math. Stat.*, 39(2):347–357, 1968.
- [3] K. B. Athreya and P. E. Ney. *Branching Processes*, pages 199–206. Springer-Verlag, Berlin, 1972.
- [4] E. Clayton, D. P. Dou   , A. M. Klein, D. J. Winton, B. D. Simons, and P. H. Jones. A single type of progenitor cell maintains normal epidermis. *Nature*, 446(7132):185–189, 2007.
- [5] D. R. Cox and H. D. Miller. *The Theory of Stochastic Processes*, pages 257–262. Methuen & Co. Ltd., London, 1965.
- [6] R. V. dos Santos and L. M. da Silva. The noise and the KISS in the cancer stem cells niche. *J. Theor. Biol.*, 335:79–87, 2013.
- [7] R. Durrett. *Probability: Theory and Examples*, pages 58–59. Cambridge University Press, Cambridge, 4th edition, 2010.
- [8] I. J. Fidler and M. L. Kripke. Metastasis results from preexisting variant cells within a malignant tumor. *Science*, 197(4306):893–895, 1977.
- [9] C. M. Fillmore and C. Kuperwasser. Human breast cancer cell lines contain stem-like cells that self-renew, give rise to phenotypically diverse progeny and survive chemotherapy. *Breast Cancer Res.*, 10(2):R25, 2008.
- [10] A. Golubev. Exponentially modified gaussian (EMG) relevance to distributions related to cell proliferation and differentiation. *J. Theor. Biol.*, 262:257–266, 2010.

- [11] P. B. Gupta, C. M. Fillmore, G. Jiang, S. D. Shapira, K. Tao, C. Kuperwasser, and E. S. Lander. Stochastic state transitions give rise to phenotypic equilibrium in populations of cancer cells. *Cell*, 146(4):633–644, 2011.
- [12] E. D. Hawkins, M. L. Turner, M. R. Dowling, C. van Gend, and P. D. Hodgkin. A model of immune regulation as a consequence of randomized lymphocyte division and death times. *Proc. Natl. Acad. Sci.*, 104:5032–5037, 2007.
- [13] P. Jagers. The proportions of individuals of different kinds in two-type populations. A branching process problem arising in biology. *J. Appl. Probab.*, pages 249–260, 1969.
- [14] S. Janson. Functional limit theorems for multitype branching processes and generalized Pólya urns. *Stoch. Process. Appl.*, 110(2):177–245, 2004.
- [15] S. Karlin and H. M. Taylor. *A First Course in Stochastic Processes*, pages 547–551. Academic Press, New York, 2nd edition, 1975.
- [16] H. Kesten and B. P. Stigum. Limit theorems for decomposable multi-dimensional Galton-Watson processes. *J. Math. Anal. Appl.*, 17(2):309–338, 1967.
- [17] M. Kimmel and D. E. Axelrod. *Branching Processes in Biology*, pages 103–140. Springer, New York, 2002.
- [18] E. Kussell and S. Leibler. Phenotypic diversity, population growth, and information in fluctuating environments. *Science*, 309(5743):2075–2078, 2005.
- [19] C. J. Mode. *Multitype Branching Processes: Theory and Applications*, volume 34, pages 138–145. American Elsevier Pub. Co., New York, 1971.
- [20] J. D. Murray. *Mathematical Biology I: An Introduction*, pages 1–6, 101–105. Springer-Verlag, Berlin, Heidelberg, 2001.
- [21] J. R. Norris. *Markov Chains*, page 122. Cambridge University Press, Cambridge, 1997.
- [22] E. M. Ozbudak, M. Thattai, H. N. Lim, B. I. Shraiman, and A. van Oudenaarden. Multistability in the lactose utilization network of *Escherichia coli*. *Nature*, 427(6976):737–740, 2004.
- [23] L. Patrawala, T. Calhoun, R. Schneider-Broussard, J. Zhou, K. Claypool, and D. G. Tang. Side population is enriched in tumorigenic, stem-like cancer cells, whereas ABCG2<sup>+</sup> and ABCG2<sup>−</sup> cancer cells are similarly tumorigenic. *Cancer Res.*, 65(14):6207–6219, 2005.
- [24] E. Seneta. *Non-negative Matrices and Markov Chains*, page 22. Springer, New York, 2nd edition, 1981.
- [25] R. T. Smythe. Central limit theorems for urn models. *Stoch. Process. Appl.*, 65(1):115–137, 1996.

- [26] A. Y. Yakovlev, M. Mayer-Proschel, and M. Noble. A stochastic model of brain cell differentiation in tissue culture. *J. Math. Biol.*, 37(1):49–60, 1998.
- [27] A. Y. Yakovlev and N. M. Yanev. Relative frequencies in multitype branching processes. *Ann. Appl. Probab.*, 19(1):1–14, 2009.
- [28] A. Y. Yakovlev and N. M. Yanev. Limiting distributions for multitype branching processes. *Stoch. Anal. Appl.*, 28(6):1040–1060, 2010.
- [29] G. Yang, Y. Quan, W. Wang, Q. Fu, J. Wu, T. Mei, J. Li, Y. Tang, C. Luo, Q. Ouyang, S. Chen, L. Wu, T. Hei, and Y. Wang. Dynamic equilibrium between cancer stem cells and non-stem cancer cells in human SW620 and MCF-7 cancer cell populations. *Br. J. Cancer*, 106(9):1512–1519, 2012.
- [30] D. Zhou, Y. Wang, and B. Wu. A multi-phenotypic cancer model with cell plasticity. *J. Theor. Biol.*, 357:35–45, 2014.
- [31] D. Zhou, D. Wu, Z. Li, M. P. Qian, and M. Q. Zhang. Population dynamics of cancer cells with cell state conversions. *Quant. Biol.*, 1(3):201–208, 2013.
- [32] M. H. Zwietering, I. Jongenburger, F. M. Rombouts, and K. van’t Riet. Modeling of the bacterial growth curve. *Appl. Environ. Microbiol.*, 56(6):1875–1881, 1990.